

# Optimal Gaussian Filter for Effective Noise Filtering

Sunil Kopparapu and M Satish

## Abstract

In this paper we show that the knowledge of noise statistics contaminating a signal can be effectively used to choose an optimal Gaussian filter to eliminate noise. Very specifically, we show that the additive white Gaussian noise (AWGN) contaminating a signal can be filtered best by using a Gaussian filter of specific characteristics. The design of the Gaussian filter bears relationship with the noise statistics and also some basic information about the signal. We first derive a relationship between the properties of the Gaussian filter, noise statistics and the signal and later show through experiments that this relationship can be used effectively to identify the optimal Gaussian filter that can effectively filter noise.

## Index Terms

Filtering, Gaussian Smoothing, Noise removal

## I. INTRODUCTION

Signal smoothing or noise filtering or denoising has been an area of active research and continues to hold the attention of researchers in various fields, for example, [1], [2], [3], [4], [5]. Noise is inherent in signals [6], [7] and a necessary first step is noise removal before any other processing can take place. A successful pre-processing step to remove noise improves the performance of the *actual processing* on the signal [8]. There are essentially two ways of taking care of noise in the signal, namely, (a) pre-processing of the signal to enable noise removal or (b) use of a set of robust algorithms that can compensate for the inherent noise. In signal processing literature pre-processing of the signal is the preferred approach.

### A. Problem

Let  $X = [x_1, x_2, \dots, x_N]$  be a band limited ( $\mathcal{B}$ ) digitized signal which is sampled at a sampling frequency of  $f_s$  and Let  $N = [n_1, n_2, \dots, n_N]$  be the noise sequence. Further assume that  $\{n_i\}_{i=1}^N$  is Gaussian distributed with mean  $\mu_N$  and variance  $\sigma_N^2$ . Let

$$X_N = X + N \quad (1)$$

represent the signal  $X$  contaminated by AWGN  $N$ . Now the problem can be stated as, given  $X_N$  estimate  $\hat{X}$  such that the error in the estimate is minimum, namely

$$\min_{\hat{X}} \|X - \hat{X}\|^2 \quad (2)$$

Typically the process of estimating  $\hat{X}$  given the noise contaminated  $X_N$  is called noise filtering or denoising. We will restrict our discussion, in this paper to the usage of a Gaussian smoothing filter for noise removal. We describe Gaussian filtering in Section II which is characterized by  $\sigma_f$  which determines the amount of smoothing. We build theory in Section III which

allows identification of an optimal  $\sigma_f^{opt}$ . We show experimentally how the identification of the actual Gaussian filter can be found in Section IV and conclude in Section V.

## II. GAUSSIAN SMOOTHING

A Gaussian filter is parametrized by its means  $\mu_f$  and variance  $\sigma_f^2$  and represented by

$$G_f(\mu_f, \sigma_f^2, t) = \frac{1}{\sqrt{2\pi\sigma_f^2}} \exp\left\{-\frac{(t-\mu_f)^2}{2\sigma_f^2}\right\} \quad (3)$$

*Note 1:* Given  $\mu_f$  and  $\sigma_f^2$  one can construct a Gaussian filter (3) with  $t$  running between  $[-\infty, \infty]$ .

*Note 2:* It is well known that spanning  $t$  between  $-3\sigma_f$  and  $3\sigma_f$  covers 99.7 % of the total area under the Gaussian.

So we can approximate  $G_f(\mu_f, \sigma_f^2, t)$  from  $t = -\infty$  to  $\infty$  as  $G_f(\mu_f, \sigma_f^2, t)$  from  $t = -3\sigma_f$  to  $3\sigma_f$  for the purpose of discussion and subsequent experimentation. Let the discrete version of  $G_f(\mu_f, \sigma_f^2, t)$  from  $t = -3\sigma_f$  to  $3\sigma_f$  be represented by  $G_f[\mu_f, \sigma_f^2, m]$  from  $m = -\lceil 3\sigma_f \rceil$  to  $\lceil 3\sigma_f \rceil$ , where  $\lceil \bullet \rceil$  represents the ceil of  $\bullet$ . Let  $X_N$  smoothed with  $G_f[\mu_f, \sigma_f^2, \cdot]$  result in  $\hat{X}^{\sigma_f^2}$ , namely,

$$\begin{aligned} \hat{X}_k^{\sigma_f^2} &= \sum_{i=-\lceil 3\sigma_f^2 \rceil + k}^{\lceil 3\sigma_f^2 \rceil + k} X_{N_i} G_f[\mu_f, \sigma_f^2, i - k] \\ &= \sum_{i=-\lceil 3\sigma_f^2 \rceil + k}^{3\lceil \sigma_f^2 \rceil + k} (x_i + n_i) G_f[\mu_f, \sigma_f^2, i - k] \end{aligned} \quad (4)$$

for  $k = 1, 2, \dots, N$ . Let the error in the estimate be

$$E_{\sigma_f^2} = \frac{1}{N} \sum_{k=1}^N \left( X_k - \hat{X}_k^{\sigma_f^2} \right)^2 \quad (5)$$

We hypothesize that one can achieve an optimal estimate  $\hat{X}_k^{\sigma_f^2}$  for some  $\sigma_f^2$  such that  $E_{\sigma_f^2}$  is minimized. We further hypothesize that  $\sigma_f^2$  is based on the variance of the noise affecting the signal and some properties of the signal. Specifically,  $\sigma_f^2$  is dependent directly or indirectly on  $\sigma_N^2$  and  $\mathcal{B}$ .

## III. OUR APPROACH

In the frequency domain we can write (1) as

$$X_N(\omega) = X(\omega) + N(\omega) \quad (6)$$

and the Gaussian filter as

$$G(\omega) = \exp\left(\frac{-\omega^2 \sigma_f^2}{2}\right) \quad (7)$$

The estimate of the signal  $\hat{X}_N(\omega)$  due to filtering by Gaussian filter can be written as

$$\hat{X}_N(\omega) = X(\omega)G(\omega) + N(\omega)G(\omega) \quad (8)$$

The error in the filtered output is given by

$$\begin{aligned} E(\omega) &= X(\omega) - \hat{X}_N(\omega) \\ &= \underbrace{X(\omega)[1 - G(\omega)]}_{\text{Signal Distortion}} + \underbrace{N(\omega)G(\omega)}_{\text{Noise Smoothing}} \end{aligned} \quad (9)$$

As seen in (9) the error in the estimate ( $E(\omega)$ ) due to filtering has two components namely, one due to distortion of signal ( $X(\omega)[1 - G(\omega)]$ ) and the other due to the reminiscent noise ( $N(\omega)G(\omega)$ ) in the signal after filtering. Let  $P_\bullet$  denote the power in the signal  $\bullet$ , then input and output signal to noise ( $\mathcal{S}$ ) ratios are given by

$$\begin{aligned} \mathcal{S}_i &= \frac{P_X}{P_N} \\ \mathcal{S}_o &= \frac{P_X}{P_X - P_{\hat{X}}} = \frac{P_X}{P_E} \end{aligned} \quad (10)$$

*Note 3:* For a certain  $\sigma_f^2$ , the Gaussian filter is able to filter the signal such that  $\mathcal{S}_o > \mathcal{S}_i$ . Namely, simultaneously remove the noise and not distort the signal.

*Note 4:* If we increase  $\sigma_f^2$  then the cutoff frequency and the bandwidth of Gaussian filter will decrease as seen in (7) and subsequently this will lead to more noise removal but on same account the signal distortion will also increase.

In the limiting case when  $\sigma_f \rightarrow 0$ , we have an all pass filter and hence  $\mathcal{S}_o = \mathcal{S}_i$ . Let for some  $\sigma_f^2 = \sigma_{fR}^2$   $\mathcal{S}_o = \mathcal{S}_i$ , such that if we increase  $\sigma_f^2$  further then  $\mathcal{S}_o < \mathcal{S}_i$ . One can hypothesize that for  $\sigma_f^2$  in the range  $[0, \sigma_{fR}^2]$ ,  $\mathcal{S}_o > \mathcal{S}_i$ . We further hypothesize that there exists a  $\sigma_{f,opt}^2$  (in the range  $[0, \sigma_{fR}^2]$ ) for which  $\mathcal{S}_o$  peaks to achieve  $\mathcal{S}_o^{max}$ . We show through curve fitting and later experimentally that we can determine the optimal  $\sigma_{f,opt}^2$  such that  $\mathcal{S}_o$  is maximized.

#### A. Determining $\sigma_{f,opt}^2$

With an aim to identify  $\sigma_{f,opt}^2$  the optimal choice of Gaussian filter to remove noise we constructed three different signals ( $X$ ) with different bandwidths ( $\mathcal{B}$ ). We constructed the noisy signal ( $X_N$ ) by appending  $X$  with  $N$  with varying  $\sigma_N^2$ . For each of this noisy signal we used different  $\sigma_f^2$  Gaussian to filter noise and for each of this  $\mathcal{S}_o$  is computed. The band limited  $X$  is constructed by first generating a random sequence of length  $\mathcal{N}$  having a normal distribution with mean zero and variance one. This random signal is smoothened using a filter of length  $\mathcal{M}(\ll \mathcal{N})$ . The impulse response of the smoothing filter is given by

$$\begin{aligned} h(m) &= 1 \text{ for } 0 \leq m \leq \mathcal{M} - 1 \\ &= 0 \text{ otherwise} \end{aligned} \quad (11)$$

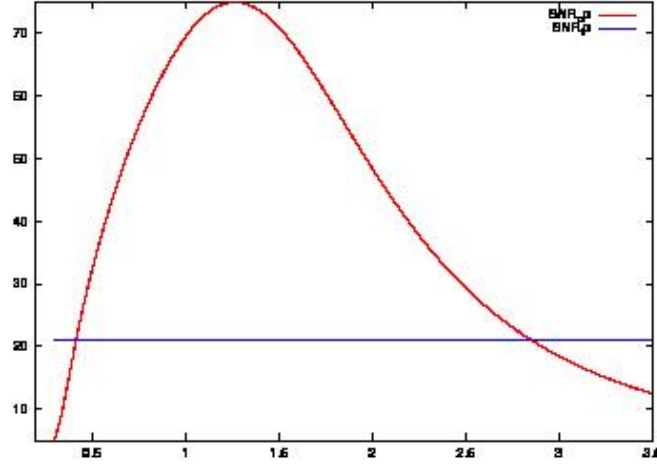


Fig. 1. The  $\mathcal{S}_o$  of filtered  $X_N^2$  for different values of  $\sigma_f^2$ .

Note that if we take a  $\mathcal{N}$  point DFT of this smoothed signal, then most of the energy is limited to  $f_s/\mathcal{M}$  Hz or  $\mathcal{N}/\mathcal{M}$  points. We cut off the high frequency region of the signal, namely, we set the points from  $\mathcal{N}/\mathcal{M}$  to  $(\mathcal{N} - \mathcal{N}/\mathcal{M})$  to zero. The inverse DFT of this low-pass filtered signal is the test signal with maximum frequency  $f_{max} = f_s/\mathcal{M}$  Hz. Note that different values of  $\mathcal{M}$  produce a filtered signal with different  $f_{max}$  and hence bandwidths ( $\mathcal{B}$ ). In this manner we constructed three different signals, each of length  $\mathcal{N} = 1024$  with  $\mathcal{M} = 5, 7, 10$ . We denote these three signals as  $X_5, X_7$  and  $X_{10}$  having  $f_{max}$  of  $\frac{f_s}{5}, \frac{f_s}{7}, \frac{f_s}{10}$  Hz respectively. An additive white Gaussian noise with  $\sigma_N^2 = 30, 35$  and  $40$  denoted by  $N_{30}, N_{35}, N_{40}$  is generated. In all we had 9  $X_N$  as our test bed. Namely,  $X_N^1 = X_5 + N_{30}$ ,  $X_N^2 = X_5 + N_{35}$ ,  $X_N^3 = X_5 + N_{40}$ ,  $X_N^4 = X_7 + N_{30}$ ,  $X_N^5 = X_7 + N_{35}$ ,  $X_N^6 = X_7 + N_{40}$ ,  $X_N^7 = X_{10} + N_{30}$ ,  $X_N^8 = X_{10} + N_{35}$ ,  $X_N^9 = X_{10} + N_{40}$ .

These signals  $\{X_N^k\}_{k=1}^9$  are denoised using a Gaussian filter (3) with different  $\sigma_f^2$ . We varied  $\sigma_f^2$  from 0.3 to 3.5 in steps of 0.01 (320 data points). For all these filtered output signal, namely,  $\hat{X}$ , the  $\mathcal{S}_o$  is calculated. Fig. 1 shows the  $\mathcal{S}_o$  of the filtered  $X_N^2$  for different values of  $\sigma_f^2$ . The x-axis shows the different values of  $\sigma_f^2$  and the bell shaped curve is the  $\mathcal{S}_o$ ; also  $\mathcal{S}_i$  (23 dB) is shown as a horizontal line. We had 320  $\mathcal{S}_o$  for varying  $\sigma_f^2$  for each of the 9 noisy signals. We now try to fit a curve so as to relate the  $\mathcal{S}_o$  in terms of  $\mathcal{B}$ ,  $\mathcal{S}_i$  and  $\sigma_f^2$ . We did this in two steps using [9].

Step 1 For a fixed  $\mathcal{B}$ , we fit a 3-D curve to relate  $\mathcal{S}_o$ ,  $\mathcal{S}_i$  and  $\sigma_f^2$  for  $\mathcal{B} = 5, 7, 10$  separately using the reciprocal full quadratic function<sup>1</sup>, namely,

$$\begin{aligned} \mathcal{S}_o = & \{a_{\mathcal{B}} + b_{\mathcal{B}}\sigma_f + c_{\mathcal{B}}\mathcal{S}_i + d_{\mathcal{B}}\sigma_f^2 \\ & + f_{\mathcal{B}}\mathcal{S}_i^2 + g_{\mathcal{B}}\sigma_f\mathcal{S}_i\}^{-1} \end{aligned} \quad (12)$$

with minimize the sum of squared absolute error criteria. For each  $\mathcal{B} = 5, 7, 10$  we obtained a set of coefficients  $a, b, c, d, f$  and  $g$ , so in all we had 18 coefficients, namely,  $A = [a_5, a_7, a_{10}]$ ,  $B = [b_5, b_7, b_{10}]$ ,  $C = [c_5, c_7, c_{10}]$ ,  $D = [d_5, d_7, d_{10}]$ ,  $F = [f_5, f_7, f_{10}]$  and  $G = [g_5, g_7, g_{10}]$

Step 2 We then fit a quadratic curve for each coefficient set, namely,  $A, B, C, D, F, G$  and  $\mathcal{B}$  separately. Using  $A$  we found

<sup>1</sup>Experimented with several functions before converging onto the reciprocal full quadratic function

$\mathcal{B}$	$\sigma_N$	$\mathcal{S}_i$	$\sigma_{f,opt}$ (14)	$\sigma_{f,opt}$	$\mathcal{S}_o^{max}$ (15)	$\mathcal{S}_o^{max}$
10	30	28.6	1.23	1.18	96.2	95.5
10	35	21.0	1.33	1.26	74.5	75.0
10	40	16.1	1.39	1.34	59.8	60.9
7	30	39.2	0.90	0.87	95.8	96.3
7	35	28.8	0.97	0.92	74.3	75.6
7	40	22.0	1.00	0.98	59.9	61.3
5	30	52.2	0.65	0.65	90.2	91.4
5	35	38.4	0.69	0.69	70.5	72.2
5	40	29.4	0.72	0.72	57.2	58.9

TABLE I  
COMPARISON OF ACTUAL  $\sigma_f$ ,  $\mathcal{S}_o^{max}$  WITH DERIVED  $\sigma_f$  USING (14),  $\mathcal{S}_o^{max}$  USING (15).

that  $a$  in (12) is related to  $\mathcal{B}$  as  $a = \alpha_1 + \alpha_2\mathcal{B} + \alpha_3\mathcal{B}^2$ . Similarly coefficients  $b, c, d, f, g$  can be written in terms of  $\mathcal{B}$ . Namely,

$$\begin{aligned}
a_{\mathcal{B}} &= (0.8364 - 1.504\mathcal{B} + 4.017\mathcal{B}^2) \times 10^{-1} \\
b_{\mathcal{B}} &= (-0.1790 - 2.572\mathcal{B} - 4.164\mathcal{B}^2) \times 10^{-1} \\
c_{\mathcal{B}} &= (-0.4596 + 3.313\mathcal{B} - 7.653\mathcal{B}^2) \times 10^{-2} \\
d_{\mathcal{B}} &= (0.7983 - 8.658\mathcal{B} + 1.575\mathcal{B}^2) \times 10^{-2} \\
f_{\mathcal{B}} &= (0.7481 - 6.817\mathcal{B} + 17.04\mathcal{B}^2) \times 10^{-4} \\
g_{\mathcal{B}} &= (0.5562 - 2.510\mathcal{B} + 6.352\mathcal{B}^2) \times 10^{-3}
\end{aligned} \tag{13}$$

Now we have (12), we get  $\sigma_{f,opt}$  by differentiating (12) with respect to  $\sigma_f$  and setting

$$\frac{\partial \mathcal{S}_o}{\partial \sigma_f} = 0$$

namely,

$$\sigma_{f,opt} = -\frac{g_{\mathcal{B}}\mathcal{S}_i + b_{\mathcal{B}}}{2d_{\mathcal{B}}} \tag{14}$$

where  $g_{\mathcal{B}}, b_{\mathcal{B}}, d_{\mathcal{B}}$  are given in (13). We get  $\mathcal{S}_o^{max}$  by substituting the value of  $\sigma_{f,opt}$  in (12), namely,

$$\begin{aligned}
\mathcal{S}_o^{max} &= \{a_{\mathcal{B}} + b_{\mathcal{B}}\sigma_{f,opt} + c_{\mathcal{B}}\mathcal{S}_i + d_{\mathcal{B}}\sigma_{f,opt}^2 \\
&\quad + f_{\mathcal{B}}\mathcal{S}_i^2 + g_{\mathcal{B}}\sigma_{f,opt}\mathcal{S}_i\}^{-1}
\end{aligned} \tag{15}$$

#### IV. EXPERIMENTAL RESULTS

We conducted a number of experiments to verify the correctness of (14) and (15) in identifying  $\sigma_{f,opt}$  and  $\mathcal{S}_o^{max}$  respectively, these results are shown in Table I and Table II. Table I tries to access the goodness of the curve fit, namely, the choice of the

$\mathcal{B}$	$\sigma_N$	$\mathcal{S}_i$	$\sigma_{f,opt}$ (14)	$\sigma_{f,opt}$	$\mathcal{S}_o^{max}$ (15)	$\mathcal{S}_o^{max}$
8	30	34.9	1.02	0.98	96.3	96.4
8	35	25.7	1.09	1.04	75.7	75.6
8	40	19.7	1.14	1.11	60.9	61.3
4	35	46.8	0.55	0.57	55.0	71.2
4	40	35.8	0.57	0.60	54.6	57.8
12	30	24.1	1.40	1.42	89.3	97.4
12	35	17.7	1.51	1.52	68.0	79.6
12	40	13.6	1.58	1.61	55.0	62.2

TABLE II  
COMPARISON OF ACTUAL  $\sigma_f$ ,  $\mathcal{S}_o^{max}$  WITH DERIVED  $\sigma_f$  USING (14),  $\mathcal{S}_o^{max}$  USING (15) FOR TEST SIGNALS.

curve and the construction of (14) and (15) from the data. As can be seen, the column four ( $\sigma_{f,opt}$  calculated from (14)) and column five (actual  $\sigma_{f,opt}$  computed from the data) are very close to each other. This is to be expected when the choice of the curve to fit the data is good. However to verify the validity of our approach to identify the  $\sigma_{f,opt}$  we conducted another set of experiments. We generated several test signals with different  $\mathcal{M}$  and  $N$  with different  $\sigma_N^2$ , such that these test signals were not part of the signals used to construct (14) using curve fitting. As can be seen in Table II, the estimation of  $\sigma_{f,opt}$  using (14) is very close to the actual  $\sigma_{f,opt}$  for all signals in Table II. As expected, a similar match is seen for  $\mathcal{S}_o^{max}$  obtained using (15) and actual  $\mathcal{S}_o^{max}$ .

## V. CONCLUSIONS

Noise removal is a mandatory pre-processing step in many signal processing applications. In this paper, we have show that it is possible to identify the optimal Gaussian filter that best filters noise, under the assumption that the noise is AWGN. The major contribution of this paper is identification of a method to obtain the optimal Gaussian filter that best filters a signal contaminated with AWGN. We have shown experimentally that the identified method works well for signals whose bandwidth and the input signal to noise ratio is know. We are in the process of verifying the validity of our approach for practical signals like speech.

## REFERENCES

- [1] D. Crisan, M. Kouritzin, and J. Xiong, "Nonlinear filtering with signal dependent observation noise," *Most*, 2008.
- [2] R. Oktem, K. Egiazarian, V. V. Lukin, N. N. Ponomarenko, and O. V. Tsymbal, "Locally adaptive DCT filtering for signal-dependent noise removal," *Eurasip Journal on Advances in Signal Processing*, 2007. [Online]. Available: <http://dx.doi.org/10.1155/2007/42472>
- [3] A. Buades, A. Silva, and B. S. Santos, "On image denoising methods," *Journal of digital imaging the official journal of the Society for Computer Applications in Radiology*, vol. 4, no. 2, pp. 1–40, 2010. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/20503063>
- [4] C. Huang, H. Wang, and B. Long, "Signal denoising based on emd," *2009 IEEE Circuits and Systems International Conference on Testing and Diagnosis*, vol. 1, no. 1, pp. 1–4, 2009. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4960873>
- [5] Y. Yang and Y. Wei, "Random interpolation average for signal denoising," *IET Signal Processing*, vol. 4, no. 6, p. 708, 2010. [Online]. Available: <http://link.aip.org/link/ISPECX/v4/i6/p708/s1&Agg=doi>
- [6] V. Bruni, B. Piccoli, and D. Vitulano, "A fast computation method for time scale signal denoising," *Signal Image and Video Processing*, vol. 3, no. 1, pp. 63–83, 2008. [Online]. Available: <http://www.springerlink.com/index/10.1007/s11760-008-0060-9>

- [7] M. L. Narayana and S. K. Kopparapu, "Effect of noise-in-speech on mfcc parameters," in *Proceedings of the 9th WSEAS international conference on signal, speech and image processing, and 9th WSEAS international conference on Multimedia, internet & video technologies*, ser. SSIP '09/MIV'09. Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS), 2009, pp. 39–43. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1946497.1946503>
- [8] V. L. Lajish, V. K. Pandey, and S. K. Kopparapu, "Knotless spline noise removal technique for improved OHCR," in *Proceedings of the 2010 International Conference on Signal and Image Processing*, 2010.
- [9] ZunZun, "Curve fitting." [Online]. Available: <http://www.zunzun.com/>